

EPA Incorporation of Five Tribes Technical Comments submitted November 12, 2019
PDI and Baseline Sampling: EPA Evaluation Report dated October 29, 2019
Portland Harbor Superfund Site

Comment	How incorporated in EPA Evaluation Report
1) Figures 1-5 through 1-14. The meaning of the shading of the grid cells in these figures is unclear. The text should explain what the shading represents and how, if at all, the shading relates to the colored dots, which we assume represent results of individual sediment grabs. In many, but not all, cases, the shading of the grid cell matches the color of the dot. Where there are multiple sediment grab samples in one grid cell, sometimes the shading of the grid cell matches the grab with the highest concentration category, while in other cases it matches the grab with the lowest concentration category.	Clarification added to Section 1.1.1. Shading of the sample points represents the concentration of the individual sample location (either SRS or SMA) while the shading of the SRS grid cell represents the concentration of the baseline sample collected within that grid cell.
2) Several figures (e.g., Figures 1-13 and 1-15) have confidentiality markings. We assume these should be removed.	Confidential stamps have been removed from all figures.
3) p. 1-2. "These results suggest that the more contaminated areas within the site have experienced a negligible decrease in surface sediment COC [contaminants of concern] concentrations." This sentence should be clarified, as it is difficult to understand how it relates to the preceding sentence, which indicates that, of the 13 sediment decision units (SDUs), eight showed nonsignificant change, four showed a significant increase, and one showed a significant decrease. Alternatively, the sentence quoted above could be deleted.	Clarification added to this sentence.
4) p. 1-4. The discussion of equivalence testing should clarify that cleanup levels (CULs) need to be achieved to demonstrate compliance with site cleanup goals. Without this statement, it is unclear whether equivalence testing could fill this role.	The following sentence was added to p. 1-4: "For the cleanup to achieve compliance with the ROD, sediment concentrations in the site will need to meet CULs; equivalence testing is a method that may be used to measure this compliance."
5) p. 1-4 and Figures 1-19x. Why was equivalence testing conducted on a site-wide scale rather than at a finer scale? Is data density insufficient for a finer-scale analysis? If so, EPA should consider providing that explanation, as EPA advocates for other types of analyses to be conducted on a fine scale in several instances throughout the report.	Equivalence testing was performed on a sitewide scale to demonstrate its functionality and capabilities as per the 2017 EPA Baseline Sampling and Long-Term Monitoring Plan. Greater density data (i.e., more samples) in both the smaller and sitewide spatial scales will

Comment	How incorporated in EPA Evaluation Report
	reduce the uncertainty in the evaluation moving forwards.
6) p. 2-2. Section 2.3 is confusing and should be restructured to more clearly present its points. The connection of this section to dioxins/furans should be made at the start, and the distinction should be drawn between the purpose of this section and Section 1.1.3.	Clarifications added to p. 2-2 including the following sentence: "In this section the discussion will focus on the dioxin/furan congeners with CULs."
7) p. 3-1 and Figures 3-3 and 3-4. The categories "dynamic equilibrium: erosional," "dynamic equilibrium: neutral," and "dynamic equilibrium: depositional" should be defined; that is, EPA should describe how an area would be classified as "dynamic equilibrium" rather than simply erosional, neutral, or depositional.	Clarifications added to p. 3-1 including the following sentence: "Each individual bathymetric survey data point was indicated as depositional (> +2.5 centimeters per year [cm/yr]), erosional (< -2.5 cm/yr), or neutral (> -2.5 cm/yr to < +2.5 cm/yr) based on the difference between the bathymetric pairs."
8) Figure 4-4. We suggest that EPA add a note similar to those in Figures 4-2 and 4-3 that CULs and RALs are not shown if they are significantly greater than measured values.	Figure revised as per comment.
9) Figures 4-4, 4-5, 4-11, 4-12, 4-13. For these figures, which compare sediment trap data between the PDI Evaluation and Remedial Investigation (RI) datasets, the X-axis warrants clarification. It is labeled Transect Location, and the values are 6 and 7. What do these numbers refer to?	Clarification was added to Section 4 including the following sentence on pgs. 4-1 and 4-2: "The RI sediment trap results are grouped and averaged based on their proximity to the PDI sediment traps, as shown on Figure 4-1 ."
10) p. 4-2. EPA states that higher sediment loading occurred during RI sampling and cites Figure 4-11. It is not clear that the figure supports this statement. It appears that loadings were similar between the RI and PDI sampling. This apparent discrepancy should be clarified.	Clarifications were added to Section 4-3.
11) p. 5-1. We request that EPA clarify whether the "subsurface SMA [sediment management area] footprints" also include surface sediment data and whether a core is considered a "hit" if a RAL is exceeded at any depth.	The following sentence was added to p. 5-1: "The 3-D sediment model includes both surface and subsurface sediment data clipped at 30 cm below mudline to distinguish the surface from the subsurface SMAs."
12) Figures 5-2 through 5-12. The purple background makes these figures difficult to read, especially the various footprints. Also, it is not clear what the legends in the top right corner of most of the figures in this series represent (and whether those 3 data are	These figures were developed as interim products to support understanding of the SMAs in the Footprint Report. They will not be revised.

Comment	How incorporated in EPA Evaluation Report
displayed in the figures). We suggest these figures be revised accordingly.	
13) p. 6-1, Section 6.1 and p. 7-1, Section 7.1. In several instances, EPA describes the surface water CULs and fish tissue target levels as “very low.” This description is confusing (low compared to what?) and implies a judgment on the ability to attain these levels that the Five Tribes are not willing to support at this time. We suggest deleting this language.	The descriptor “very low” has been removed from Sections 6-1 and 7-1.
14) Figures 6-3x and 6-9x. We recommend that the lines connecting sampling points be removed, as they seem to suggest spatial trends that may not actually be present. In addition, it would be helpful to label the Downtown Reach and Upriver Reach stations.	The lines are included to more clearly show the change in concentration or mass loading between transects. The location of the transects in the site and D/U Reach are shown on Figure 6-1. These figures will not be revised.
15) Figures 6-4x. Some of the cleanup levels appear to plot close to zero. We suggest using a log axis throughout, to clarify. In addition, the labeling for the legend runs off the edge of the page and cannot be completely read.	Figures revised as per comment.
16) Figures 6-10x. The legend should indicate the flow for Round 3, as it does for Rounds 1 and 2.	Figures revised as per comment.
17) Figures 6-16x. These figures compare a range (with average) for the PDI data with a box and whisker plot for the RI data. The text should better explain what the box and whisker plots represent. We further suggest that fewer graphs be presented per page to better allow the reader to discern areas of overlap or nonoverlap.	Clarification added to Section 6.4 to enhance the description of the box and whisker plots.
18) p. 6-4. The report states, “This evaluation in general found reductions in organic concentrations between the RI and PDI water quality sampling events but little change in metals concentrations.” It is not readily apparent that the data support this observation. We suggest deleting the statement and letting the figures speak for themselves, particularly as the text asserts that quantitative comparisons cannot be reasonably made.	This sentence has been deleted as per comment.
19) p. 7-1. The report states, “The overlap of the lipid distribution in the different years and reaches suggests that there are not statistical differences in lipids between the study years or reaches that would necessitate lipid normalization.” This statement is	The sentence in question has been clarified.

Comment	How incorporated in EPA Evaluation Report
<p>not well supported. First, distributions can overlap substantially and yet be statistically distinct. Second, visual inspection suggests that some differences are likely—e.g., the distribution of 2002 site reach data look different than the 2002 upriver reach data. The 2012 site reach data seem to have a different distribution than the 2012 downtown reach data, and the 2019 downtown and upriver reaches also look quite different from one another. If these were plotted on top of one another, these apparent differences would be more visually evident. Whether or not distributions are similar should not be determined by visual inspection. If it is desirable to compare distributions, we recommend that it be done more formally.</p>	
<p>20) p. 7-1. The report states, “Average concentrations of total PCBs, DDX, and 2,3,4,7,8-PeCDF were higher in the site than in the Downtown Reach and Upriver Reach.” This and similar quantitative comparison statements throughout should be supported with a parenthetical indicating the statistical test used to support this statement and the associated p-value (e.g., Kruskal-Wallis, $p=...$). Alternately, the text should clarify that statistical approaches were not used to evaluate differences due to [reasons].</p>	<p>Additional clarification added to Section 7.1 for the statistical comparisons of the fish tissue average concentrations.</p>
<p>21) p. 7-2 and Figures 7-17x. The report states, “there do not appear to be any substantial differences in the distributions between the proposed river segments within the site.” Whether or not distributions are similar should not be determined by visual inspection, particularly given the very long tails on these distributions. If it is desirable to compare distributions, a formal comparison should be conducted.</p>	<p>EPA does not support the use of the Pre-RD Group’s proposed river segments as they are not a ROD spatial scale and were not evaluated in the baseline risk assessments. Therefore, only visual observation was performed to describe the contaminant distributions within the proposed river segments. Statistical differences between the proposed river segments would not warrant their use for compliance with RAOs as they are not based on the exposures quantified in the baseline risk assessments.</p>
<p>22) Section 7.3. We believe that the concentration trend analysis warrants a more statistically robust approach. The statistical evaluation of time trends is a complicated, highly technical subject, and the report text does not address these complexities adequately. We recommend that the model used for time trend</p>	<p>Additional clarification added to Section 7.3. The mixed effects first order decay model was selected to generate an estimate of temporal trends.</p>

Comment	How incorporated in EPA Evaluation Report
analysis be carefully selected and the model choice explained. In addition, the potential for autocorrelation should be evaluated and taken into account.	
23) Figures 7-14 and 7-21 display error bars, but the text does not explain what the error bars represent. This information should be provided. All figures with error bars should be checked to ensure that their meaning is described in the text.	A description of the confidence intervals was added to Sections 7.1 and 7.3.
24) Section 9 uses qualitative terms to describe fish movement (e.g., “exhibited little movement,” “stationary”). It is difficult to determine how these qualitative descriptions compare to the RI assumption of a 1-mile home range. We suggest that EPA describe fish movement in quantitative terms where possible.	The Figure 9-4 series shows the estimated locations of the smallmouth bass during the fish tracking study. Home ranges were not calculated due to the limitations of the study design, discussed in Section 9.2.
25) Section 10. The analysis supporting the selection of a smaller sediment average radius (of 100 to 600 feet) could be strengthened. In particular, the text identifies the “strongest relationships” based on an analysis of p-values as a function of sediment averaging radius (results shown in Figure 10-1). Even ignoring the fact that 0.05 is an arbitrary, although frequently used, p-value for statistical significance, the results for most contaminants suggest that a radius of up to the maximum for which data are presented (around 1300 feet) seem to produce approximately equally good results—certainly the p-values are below 0.05 for four of the five evaluated COCs, for radii well beyond the 600-foot mark. We do not believe that this analysis can differentiate between different hypothesized averaging radii. We suggest consultation with a statistician and spatial modeler to identify alternate approaches to address this question. Two possibilities may be worth exploring. For instance, if alternative models using different radii were developed, they could potentially be compared using the Akaike information criterion. In addition, modeling approaches are available to identify the size and locations of core use areas (e.g., see Hart et al. 2010, Hart et al. 2013). ²	EPA disagrees with the assessment that this analysis cannot differentiate between different averaging radii. For a fixed number of regression variables, the minimum p-value and minimum Akaike information criterion will identify the same model. Additionally, the modeling approaches identified in Hart et al. 2010 and Hart et al. 2013 (kernel density estimates and maximum convex polygons) are for telemetered specimens and are unrelated to the regression analysis performed in Section 10. The regression analyses performed in this section were developed to refute the claim from the Pre-RD Group that there was no spatial relationship between concentrations of bioaccumulative organic COCs in sediment and fish tissue, not to replace the mechanistic food web model that underpins the RAO 2 sediment CULs and fish tissue target levels. The regression analysis was performed by Dr. John Kern, a sediment geostatistician with more than 20 years of experience.